Available online at www.rpnsci.com

ISSN (P): 3006-8452, ISSN (O): 3006-8460

Research Prospects in Natural Sciences, Vol. 2, Issue 2, (2024) 1–12

Review Article

Role of hepatitis C virus envelope proteins in vaccine design

Areeba Maryam[a,*], Komal Fatima[a], Areeba But[a]

[a]*Department of Biomedical Engineering, University of Engineering and Technology Lahore, Narowal Campus, Pakistan*

Abstract

Hepatitis C virus (HCV) infection remains a significant global health problem, affecting millions of individuals with confirmed liver conditions. Despite recent advancements in antiviral therapies, the development of an effective Vaccine remains a top priority. Understanding the molecular epidemiology and viral evolution of hepatitis C virus (HCV) infection requires sequencing. Due in part to the tremendous genomic variety found in HCV, there is currently limited uniformity among sequencing methodologies. HCV envelope proteins, in particular E1 and E2, are critical targets for vaccine development because they play a critical role in the infection's entrance into host cells. These envelope proteins, particularly E1 and E2, are pivotal in facilitating the original intercourse between the virus and the host's vulnerable system, making them ideal targets for vaccine design. E1 and E2 are complex molecules with different structures, featuring glycosylation sites and containing essential antigenic regions that are critical for the host's vulnerable response. Still, their structural diversity poses a significant challenge in vaccine development. HCV is well- known for its inheritable variability, with multiple genotypes and quasi species circulating globally. This inheritable diversity has implications for the effectiveness of any potential vaccine, as a single strain may not give sufficient protection against the wide range of HCV variants. To address this challenge, innovative strategies are being explored. A general time-reversible substitution model was used to build phylogenetic trees, and sensitivity studies were carried out for each subregion.

Keywords:
Hepatitis C Virus, Antiviral Therapies, HCV Vaccine, Antigenic Regions, Conserved Epitopes

1. Introduction

Hepatitis C virus (HCV) was initially identified by Michael Houghton and associates in 1989 through cloning and sequencing of HCV genome collected from infected chimpanzees [1]. The virus is responsible for hepatitis C, which is an important global complaint chronically infecting around 57.8 million people. Hepatitis C virus significantly increases the incidence of health complications and fatality associated with liver diseases by elevating the chances of developing conditions such as liver cancer, cirrhosis, and liver fibrosis. Hepatitis C accounts for over 399,000 fatalities worldwide each year [2]. Patients typically acquire HCV through blood contact. Therefore, the highest risk groups consist of intravenous drug users, people who experienced blood transfusions before 1992, and healthcare professionals. Sexual transmission is now a significant problem among MSM who are HIV-positive, regardless of the fact that it is responsible for a very minor portion of the causes of acute hepatitis C. Seven to twenty-one days after the virus has been transmitted, HCV RNA can be found in serum. But prolonged incubation times can happen, particularly when only little virus loads have been transmitted [3].

The HCV symptoms starts to appear 3 to 12 weeks after exposure. The signs include fatigue, appetite loss, and jaundice. Serum alanine aminotransferase (ALT) levels rise within two to eight weeks of exposure, frequently exceeding the upper limits of the normal range. A week or two after exposure, serum can include HCV RNA. In the early weeks of infection, HCV RNA levels rise rapidly, reaching a peak between 105 and 107 IU/mL before the elevation of serum aminotransferase levels and the appearance of symptoms [4]. The majority of individuals with acute HCV infection approximately 80% to 85% fail to clear the virus and progress to chronic infection. Long-term persistence of the virus can result in serious complications, including cirrhosis, portal hypertension, hepatic decompensation with encephalopathy, and hepatocellular carcinoma [5].

2. Global Epidemiology

According to the World Health Organization (WHO), there are approximately 1.75 million new hepatitis C virus (HCV) in-

*Corresponding Author:
2020bme136@student.uet.edu.pk (Areeba Maryam)

fections annually, equating to a global incidence rate of 23.7 cases per 100,000 people. The Centers for Disease Control and Prevention (CDC) reported that hepatitis C-related deaths increased fourfold in 2020 compared to 2019. HCV prevalence varies significantly across different regions, ranging from 0.5% to 6.5%. In India, the reported prevalence is 0.9%, while it ranges from 0.5% to 1.5% in Western countries. Other regional estimates include 2.2% in Indonesia, 2.3% in Southeast Asia and Eastern Mediterranean regions, 6.5% in Pakistan, and 3.2% in China. WHO also estimates that 10 million people are living with chronic hepatitis C in the Western Pacific and Southeast Asia. Among Asian regions, Southeast Asia ranks just behind East Asia and South Asia in terms of HCV-related deaths [1].

The highest incidence is around 17.5% in Africa, 55.3% in Asia, 93.1% in the US, 88.1% in the Pacific, and 75.1% in Europe (Figure 1) [6]. Overall, it was projected that in 2019, there were 6.2 million new HCV infections, 540,000 HCVrelated deaths, and 15.3 million DALYs (disability-adjusted life years) associated with HCV, these figures represent increases of 25.4,59.1, and 43.6, respectively, from 1990. Men and women both have equal numbers of new HCV infections globally in 2019, but men endured more deaths and disability- adjusted life years (DALYs), as well as a bigger increase, than women. India and China were the two countries where nearly one-sixth of all new infections around the world started. India had the most DALYs connected to HCV infection among the three countries (India, China, and the United States), despite China having the most HCV infection-related mortality. Men had higher rates of incidence, fatalities, and DALYs than women in all three nations [7].

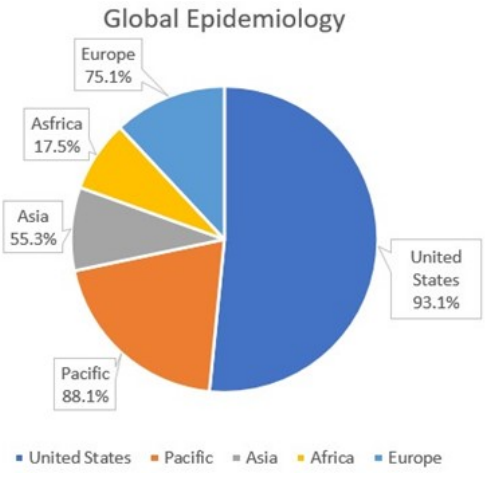

Figure 1: The highest frequency is around 17.5% in Africa 55.3% in, Asia , the 93.1% in US, 88.1% in the pacific , and 75.1% in Europe

In Pakistan the epidemiology of the hepatitis C virus was 16.47% in the general population, this study includes 765,426 individuals. While 8.2% in blood benefactors which covers the total 973,260 individuals. The average frequency in pregnant women (136,546) was 9.3%. This demonstrates that the prevalence of Hepatitis C was 11.32 percent in the general population, pregnant women, and blood benefactors [8].

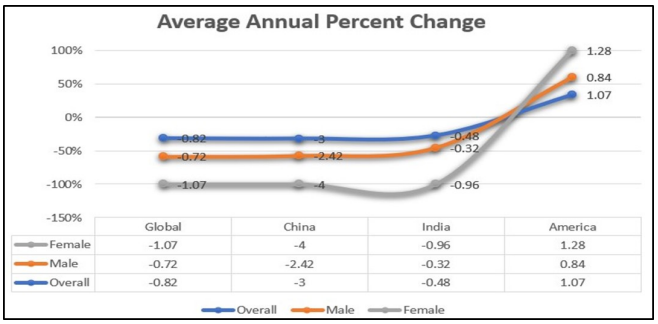

Figure 2: Age-standardized mortality rate (ASMR), average annual percent change (AAPC) from 1990 to 2019

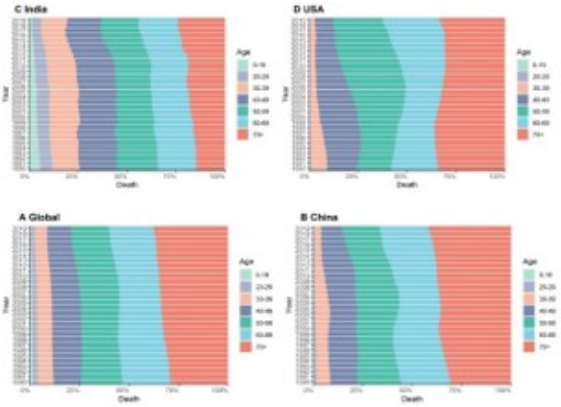

Figure 3: Age-specific HCV-related death rates by area from 1990 to 2019 (a) Global (b) China (c) India (d) The United States [7]

The age standardized incidence rate of the HCV infection (ASIR) worldwide in 2019 was 82.5 per 100,000 people, with a 95% confidence interval spanning 73.4 to 94.7. With a 95% UI range of 5.9 to 7.5, the age standardized mortality rate (ASMR) was 6.7 per 100,000 people. The age standardized delay rate (ASDR), with a 95% UI range from 160.7 to 210.2, was 184.5 per 100,000 people.

The average annual percent changes (AAPCs) for these three measures were, respectively, -0.30, -0.82, and -0.95 from 1990 to 2019. India's ASDR (Age Standardized Death Rate) was comparable to the global average. Age Standardized Incidence Rate (ASIR) was greatest in the United States and lowest in China. The Age Standardized Mortality Rate (ASMR) and Age Standardized Death Rate (ASDR) of the United States had the largest rise ratio among the three countries. For the United States, these values were 1.28 and 1.30, respectively. On the other side, China has seen a dramatic decline in the disease burden of HCV, especially among women. Among the three nations, Chinese women saw the biggest drops in ASDR and ASMR for HCV infection.. The corresponding AAPCs for Chinese women were -4.00 and 4.35, respectively [7].

People over the age of 70 accounted for more than 30% of HCV-related deaths worldwide, in China, and in the US in

2019. Among them, the United States' expanding population mortality rate is more pronounced, with a higher percentage of people in the 50 to 69 age group than in the rest of the globe, China, and India, as shown in (Figure 3).In India, the age distribution of the death rate is sufficiently balanced. Therefore, those who were 70 years of age or older suffered the greatest number of Hepatitis C virus-related deaths worldwide. The average annual percent changes (AAPCs) for these three measures were, respectively, -0.30, -0.82, and -0.95 from 1990 to 2019. India's ASDR (Age-Standardized Death Rate) was comparable to the global average. The Age-Standardized Incidence Rate (ASIR) was greatest in the United States and lowest in China. The Age Standardized Mortality Rate (ASMR) and Age Standardized Death Rate (ASDR) of the United States had the largest rise ratio among the three countries. For the United States, these values were 1.28 and 1.30, respectively. On the other side, China has seen a dramatic decline in the disease burden of HCV, especially among women. Among the three nations, Chinese women saw the biggest drops in ASDR and ASMR for HCV infection. The corresponding AAPCs for Chinese women were -4.00 and 4.35, respectively. (Yang, Qi et al. 2023). People over the age of 70 accounted for more than 30% of HCV-related deaths worldwide, in China, and in the US in 2019. Among them, the United States' expanding population mortality rate is more pronounced, with a higher percentage of people in the 50 to 69 age group than in the rest of the globe, China, and India, as shown in (Figure 3).In India, the age distribution of the death rate is sufficiently balanced. Therefore, those who were 70 years of age or older suffered the greatest number of Hepatitis C virus-related deaths worldwide.

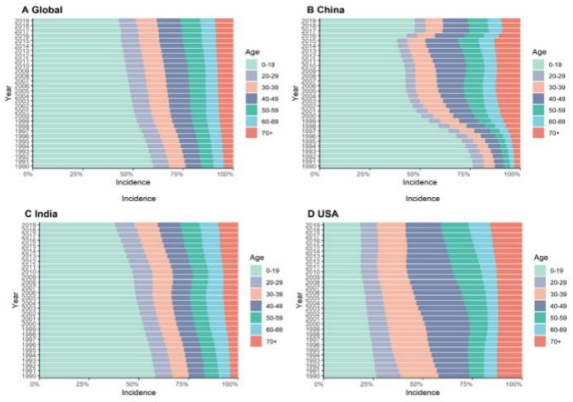

Figure 4: (A) Global (B) China (C) India (D) The United States (Constitution of incidence due to HCV by age groups in different areas from 1990 to 2019 [7]

3. Mode of Transmission

The Hepatitis C virus spreads through blood borne, perinatal, and sexual channels; however, the virus is not transferred through breastmilk, feces, or informal interactions [9].The most efficient way for HCV to spread is by prolonged or frequent direct percutaneous exposure, much like with blood transfusions and other blood products, organ transplants from contagious donors, and needle sharing among injection drug users (IDU). HCV can also spread by parenteral exposure, prenatal exposure, sexual contact, domestic contact, and parenteral exposure in a medical context. Although transmission through other unproven sources, such as contaminated instruments used in public health campaigns, traditional drugs like acupuncture, open folk drugs, tattooing, body piercing, and marketable barbering, has not been proven in the USA, similar routes may be crucial in sustaining the spread of HCV in some areas of Asia [10]. In the latter half of the 20th century, the hepatitis C virus (HCV) infection spread quickly due to the widespread availability of injectable medicines and the rise in the use of illicit injection drugs. The main risk factors for HCV transmission globally have been iatrogenic exposures and the use of illicit injectable therapies. Unsafe corrective injection procedures tend to be the leading cause of infections in underdeveloped nations. Donor testing has almost completely eliminated transfusion-related illnesses in developed nations, but infections spread to patients via improper injection procedures are a growing issue. The main risk factor for HCV is injection drug use; prevalence is still high among new injectors, and this likely explains or obscures the connections between HCV-positive people and histories of noninjecting drug use, tattooing, and incarceration. Increased use of illicit drugs may contribute to the rise in sexually transmitted HCV infections among HIV-positive men who engage in male sex [11].

4. HCV Genome

The diameter of the positive stranded, single stranded, spherical, enveloped hepatitis C virus (HCV) is between 40 and 80 nm. This virus belongs to the Flaviviridae family and is categorized under the Hepacivirus genus. The 9.6 kb HCV genome encodes a single polyprotein that is broken into ten mature proteins by cellular and viral enzymes [12]. Previous research has shown that the HCV genome is abundant in complex RNA structures, comprising useful RNA motifs in both the coding and untranslated regions (UTRs). One of the best-characterized IRES structures in any system is found within the 5´ UTR of the HCV, for instance. It is known that the core region of HCV contains highly conserved substructures, many of which are engaged in long-range interactions that are crucial for viral replication and translation [13].

With icosahedral symmetry, the nucleocapsid is made up of multiple copies of the core proteins and genomic RNA. A host cell-derived lipid bilayer envelope with the envelope glycoproteins E1 and E2 incorporated surrounds the nucleocapsid (Figure 5). Major protein components of the virion include the core protein and the E1 and E2 envelope glycoproteins. The virus's genome, which contains about 9600 nucleotides, is converted into a single precursor protein with 3020 amino acids. A large open reading frame (ORF) and highly conserved 5´ UTR and 3´ TR sections are both present in the genome. There are structural and nonstructural proteins in the HCV genome [14]. The HCV genome shows non-homogenous mutation rates throughout. It was discovered that substitution biases are largely

different between two nearby genomic regions and depend on the nucleotide composition of the genome [15].

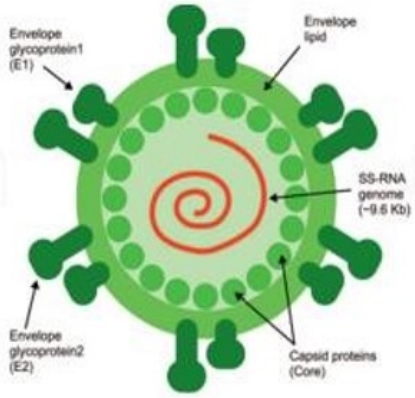

Figure 5: The Hepatitis C Virus has a single stranded, positive sense RNA molecule for its genome. It belongs to the Hepaciviral genus and the Flaviviridae family. The 9.6 kilobases (kb)-long HCV genome encodes a single polyprotein that is later degraded into separate proteins [14]. One ORF that encodes a polyprotein covers the majority of HCV genome. Viral and host proteases repurpose this polyprotein into distinct viral proteins. These proteins contain nonstructural proteins (NS1, NS2, NS3, NS4A, NS4B, NS5A, and NS5B) as well as structural proteins (Core, E1, and E2)

5. HCV Genotypes

HCV is divided into number of various genotypes and subtypes due to its genetic variability. For the accurate diagnosis, effective management, and treatment of HCV infections, it is essential to comprehend these genotypes. There are many subtypes for each of the HCV's seven main genotypes that are Genotype 1 (GT1), Genotype 2 (GT2), Genotype 3 (GT3), Genotype 4 (GT4), Genotype 5 (GT5), Genotype 6 (GT6) and Genotype 7 (GT7). It's crucial to remember that HCV genotypes can affect how the disease develops, the antiviral medications that are used, and the possibility that a given treatment will be effective. All genotypes now have considerably higher cure rates thanks to improvements in HCV therapy; however, each genotype may require a different treatment plan and time frame. The understanding of HCV genotypes is still developing, and additional genotypes and subtypes may be discovered in the future.

6. HCV Genotypes Pervalence

Understanding the distribution of HCV genotype prevalence by geographical location is crucial for developing efficient strategies for diagnosis, treatment, and prevention. There are 86 different sub-genotypes and 8 genotypes of HCV throughout the world. Because the RNA-dependent RNA polymerase (RdRp) lacks a proofreading mechanism, it is a very error-prone enzyme, leading to a significant degree of sequence variability in the freshly replicated HCV genomes. Most HCVassociated liver infections are caused by genotypes 1, 2, and 3, which are widespread worldwide and account for 90% of all infections. It is important to remember that genotypes with a global distribution are unevenly represented in various geographical areas. In the United States, Europe, Japan, and Australia, genotype 1 is more common. Genotype 3 is frequent in India, Pakistan, and Afghanistan, while genotype 2 is more common in South America, China, and Japan [16]. The most widespread genotype worldwide is GT1, which is especially prominent in North America, Europe, and Australia. While subtype 1b is more prevalent in Europe, subtype 1a is more prevalent in North America. According to a 2016 study that appeared in the Journal of Hepatology, subtype 1a of HCV was more widespread in the US and was the most common genotype among individuals with chronic infections [17]. Although GT2 is less frequent than GT1, it may still be found in many places throughout the world, including North America, Europe, and Asia. Those of subtypes 2a and 2b are more prevalent. GT2 is rather widespread in Japan. According to a 2016 study that appeared in the Journal of Gastroenterology, patients with chronic HCV were most likely to have the GT2 genotype.

South Asia, Southeast Asia, and several regions of Europe are home to GT3. GT3 is frequently linked to more severe liver disease. The most common genotype of HCV infection in India was GT3, according to a 2015 study that was published in the Journal of Medical Virology [18]. Particularly in Egypt, GT4 is widely used in the North Africa and Middle East ,the most prevalent subtype is 4a. Due mostly to GT4 infections, Egypt has among the most prevalent rates of HCV in the world. The predominance of GT4 in Egypt was described in a study that was published in the Journal of Infectious Diseases in 2008 [19]. Genotype 5 is a somewhat uncommon variant that is mostly found in a few places, with South Africa being one of these locations. Although GT5 has been found in various countries, including France, it is still primarily concentrated in South Africa. The most prevalent form of GT5 is subtype 5a [20]. Southeast Asian nations like Vietnam, Thailand, Cambodia, Laos, and Myanmar are predominantly home to the genotype 6. With varying subtype distribution, it is one of the most common genotypes in the area. The most prevalent subtype of GT6 is 6a [21]. The Hepatitis C Virus (HCV) genotype 7 (GT7) has not been generally identified or thoroughly researched. More research is required to completely understand the distribution of GT7, which is mostly prevalent in Central Africa [22].

According to recent study in Pakistan in 2023, Genotype 3 was found to be the predominant genotype, accounting for 93.75% of the samples. Other genotypes were detected at lower frequencies, with genotype 1 representing 3.25% of the samples, genotype 2 and genotype 4 each representing 1.25% of the samples. None of the samples included genotypes 5 or 6, which were undetectable. A total of 0.5% of the samples contained two recombinant Hepatitis C strains. Although one

sample could not be typed, it was found to have a genotype 3 variation. The analysis of the baseline data showed that 51.0% of the samples were of the male gender. The participants were 43 years old on average. The viral load ranged from 2x103 to 1x107 U/mL, while the mean ALTs levels were 105 U/L. In 2021, a study was conducted in Tehsil Babozi, District Swat of province Khyber Pakhtun Khaw of Pakistan, to investigate the prevalence of Hepatitis C Virus (HCV) in the population. The total population of the study area was 278,401, and a sample of 223 individuals was taken from a different site within the area [23].

The research involved both males and females to determine the gender-wise prevalence of HCV. Out of the 223 individuals analyzed, 169 were males and 53 were females. The study found that the prevalence of HCV in males was 1.78%, while in females it was 5.66%. This indicates that females were more infected with HCV compared to males. The higher prevalence of HCV in females may be attributed to factors such as lack of awareness about HCV, lower levels of education, and non-compliance with prescribed medications. Another study conducted on HCV prevalence found no statistically significant difference between males and females or between different locations. However, it revealed that there were more male patients with HCV compared to females. The highest prevalence was observed in men with a rate of 4%, while in females, the prevalence was relatively low at 1.1%. This study also indicated a higher occurrence of HCV contamination in older individuals.

The individuals included in the study ranged in age from 17 to 82, with an average age of 37 years. All patients tested positive for HCV antibodies, indicating an infection. The likelihood trends were somewhat greater among males in all age categories, but there was no statistically significant gender difference in the occurrence of HCV infection. In conclusion, women were more likely than men to have HCV in the research area. To investigate the underlying causes of this gender gap in HCV infection, more study is necessary [23]. The samples were also collected from individuals belonging to various age groups. The study included both males and females, who were divided into different age categories, such as 15-15, 16-25, 26-35, and so on, up to 105. The prevalence of HCV was found to be 1.6% in the 15-26 age group, 1.05% in the 26-35 age group, 2.7% in the 36-45 age group, 4% in the 46-55 age group, and 6.25% in the 56-65 age group. The results indicated that there is a higher prevalence of HCV in older individuals above the age of 50. This may be due to a weakened immune response to the pathogen and other agerelated diseases.

Furthermore, the results showed a statistically significant increase in the likelihood of anti-HCV prevalence with increasing age. Another study conducted on HCV prevalence found that HCV was very rare in patients aged 15 years or below, while a higher frequency of HCV was observed in older age groups. According to a survey on several age groups, the 20 to 29-year-old age range had the highest occurrence. A decline in the prevalence of active HCV was seen in age groups older than 40. According to the findings of the study, out of a total of 223 individuals, 1289 (57.4%) had no history of viral infection, while 94 (42.2%) had been affected by viral infection in the past. Among the 94 individuals with a history of viral infection, 2.12% tested positive for HCV, indicating a relationship between HCV infection and past viral infection [23]. The study also analyzed information from patients with longlasting viral hepatitis. It was discovered that patients with longlasting hepatitis B had much higher secondary, hospital, management, and overall incomes than did hepatitis B virus carriers and patients with long-lasting hepatitis C [16].

An interesting approach for the management of severe HCV and the prevention of chronic hepatitis has been proposed. Some scientific and legal actions have suggested that treating hepatitis C infection during the acute stage is associated with a high rate of sustained virological response, ranging from 75% to 100%. However, while there is a general consensus that intervention during the acute stage improves viral clearance, there are still unanswered medical questions and uncertainties regarding the optimal treatment for severe hepatitis C infection and the documentation of sustained virological response (SVR) outcomes [16]. The first significant factor in the fixed model was the patient's age, which showed a positive significance. This indicates that with an increase of 1 year in the patient's age, the risk of hepatitis C infection will increase by 1.035 times, assuming all other factors remain constant. As a one-year increase does not result in a significant change, the change was examined after 20 years. This suggested that with a 20-year increase in age, the risk of Hepatitis C disease doubles. HCV occurrence was found to be directly related to age, meaning that the older the age, the higher the occurrence of HCV [23].

7. HVC Glycoproteins

The hepatitis C virus (HCV) glycoproteins E1 and E2 are the most important targets for neutralizing antibodies. This is a direct result of their roles in mediating the virus's entry into susceptible cells, which is pH- and clathrinid-dependent. The N-terminal portion of the HCV genome houses the two genes that produce the glycoproteins. After the glycoproteins are initially made as a part of the viral polyprotein, the host cellular proteinases signal peptidase and signal peptide peptidase release the mature proteins. The mature, cleaved E1 protein has 192 amino acids, while the E2 protein has 363-369 amino acids, depending on the viral strain. The connections between the transmembrane domains of the glycoproteins, which each chaperon the folding of the other during synthesis, cause them to form heterodimers [24]. The fact that amino acid variation in the E1 and E2 proteins between infectious primary isolates exceeds 37% emphasizes the enormous genetic diversity that the E1 and E2 genes are capable of tolerating. The E2 protein's three hypervariable regions (HVRs) have the highest level of amino acid variety. The HCV polyprotein's HVR1, a region of 26-27 amino acids at the very end of E2, has the most variability. Compared to the HVR2, the intergenotypic variable region (IgVR) is situated nearer to the E2 transmembrane domain and CD81 binding regions. Despite this difference, both proteins exhibit extensive glycosylation on their surfaces as well as conserved N- and O-linked glycans. As the primary receptor binding protein, E2 interacts with the cell surface molecules CD81,

SR-BI, and occluding. While the interaction with SR-BI is assumed to be controlled by the N-terminal hypervariable region (HVR1) of the E2 protein, the binding surface between E2 and CD81 is a discontinuous surface containing three highly conserved areas of the E2 protein. It has been widely documented how E2 and cell surface receptors interact. The main result of neutralizing Abs is probably to prevent receptor interactions [25].

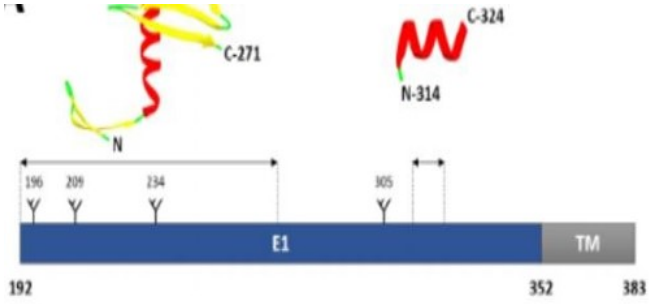

Figure 6: HCV E1 (aa192-383) in a linear diagram, along with the crystal structures of its segments aa192-271 (PDB: 4UOI) and aa314324 (PDB: 4N0Y)[25]

8. Structure of E1

The E1 glycoprotein is one of the two HCV envelope glycoproteins; the other is E2. In the early stages of HCV infection, these glycoproteins work together quite effectively. The viral envelope must fuse with the cell membrane in order for the virus to enter and infect the host cell, and E1 is essential for mediating this process. This procedure, which is also a key target for antiviral medications, is necessary for the development of HCV infection. The structure of E1 is characterized by a number of distinctive characteristics that allow it to carry out its crucial tasks during HCV infection. E1 is a type I transmembrane protein, which means that it spans and is anchored within the viral lipid envelope. The hydrophobic amino acid residues in this transmembrane region enable E1 to bind to the viral membrane. Because it is a crucial part of the viral envelope, E1 is excellently positioned for its role in aiding the fusion of the viral envelope with the host cell membrane [26].

E1 has two primary extracellular domains, E1a and E1b, which interact with host factors when the virus is infected and are visible on the viral surface. The initial attachment of the virus to the host cell surface is caused by the E1a domain. E1a has regions that interact with co-receptors and host cell receptors, making it easier for HCV to bind to the cell. The process of fusion involves E1b. Once the virus is linked to the host cell, E1b interacts with E2 and other viral and host components to induce fusion, finally allowing the viral genetic material to enter the host cell. E1 and E2 work together in the early stages of HCV infection. E1's function increases during the fusion process, whereas E2 is principally in charge of the virus's attachment to host cell receptors. The heterodimeric complex that E1 and E2 create improves their capacity to mediate viral entry.

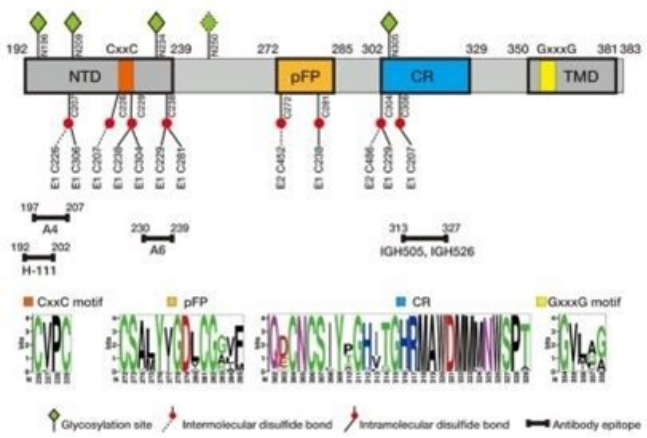

Figure 7: Hepatitis C virus (HCV) E1 envelope protein-schematic diagram [27]

The specific areas and residues within E1 and E2 that interact with each other have been the subject of intense investigation to understand how they collaborate in the infection process [24].

E1 experiences post-translational changes, like many glycoproteins. The correct folding and operation of the protein depend on these changes, which include glycosylation. E1's conformation can be impacted by glycosylation, which may also contribute to immune evasion. It's crucial to comprehend the particular glycosylation patterns of E1 in order to design vaccines and target treatments. The genetic heterogeneity of HCV is well documented, and this genetic variability extends to the E1 protein. The E1 sequences of different genotypes and subtypes of HCV differ. The effectiveness of viral entrance, the immunological response, and even the susceptibility to antiviral therapy can all be affected by this genetic variation. In order to comprehend HCV epidemiology and create individualized treatment plans, researchers explore these genetic variations in E1.

9. Functions of E1

The HCV E1 glycoprotein's actions are crucial for the virus's capacity to infect host cells and start an infection. E1 is involved in the early phases of HCV's life cycle, including host cell adhesion and membrane fusion. For the creation of antiviral treatments and vaccines to prevent HCV infections, it is essential to comprehend the roles of E1. E1a's main function is to aid in the virus's initial attachment to host cells. The first stage in establishing an HCV infection is this connection. The virus can attach to the surface of the host cell thanks to E1a's interaction with particular co-receptors and host cell receptors. E1a interacts with host cell receptors such as CD81 and class B type I scavenger receptors (SR-B1). The interaction between E1a and CD81 is essential for HCV entrance and is required for the virus to cling to host cells. Co-receptors like claudin-1 and occluding are also involved in the attachment process in addition to receptors. Viral binding to host cells is further made possible by interactions between E1a and these co-receptors [28].

Due to the genetic diversity of HCV, different genotypes and subtypes have different E1 sequences. The effectiveness of viral attachment and penetration into host cells may be impacted by this heterogeneity. The viral envelope and the membrane of the host cell are fused predominantly by E1b. For the virus to transmit its genetic material into the host cell and start the infection, this fusion process is necessary. Early on in viral entry, E1, the first major envelope glycoprotein, collaborates with E2, the second major glycoprotein. While E2 is mainly responsible for binding to receptors, E1 and E2 interact to generate a heterodimeric complex that improves the effectiveness of viral entry [29].

On going research is being done to determine how E1b mediates membrane fusion. For the purpose of creating ways to prevent viral entrance and hinder fusion, understanding this mechanism is essential. E1 is a major target in the development of vaccines because it has areas that potentially cause neutralizing antibodies. The objective is to create vaccinations that encourage the synthesis of antibodies that can neutralize HCV and stop infection [30]. Due to the virus's genetic diversity, which includes variance in E1 sequences among various genotypes and subtypes, developing an effective HCV vaccine is difficult. To guarantee widespread protection, vaccine candidates must take into account this variability [31].

10. Structure of E2

E2 protein is the second most significant glycoprotein of HCV which is characterized as type 1 transmembrane protein. It is comprised of 363 amino acids ranging from viral polyprotein position 384 to 746 [32]. It has a C terminal transmembrane domain and a N terminal ectodomain. The intensive post-translational changes that the E2 protein goes through include the creation of 9 to 11 N linked glycosylation sites and the presence of 18 cysteine residues that are conserved across all genotypes. These alterations are necessary for both immune infiltration and appropriate protein folding [33].

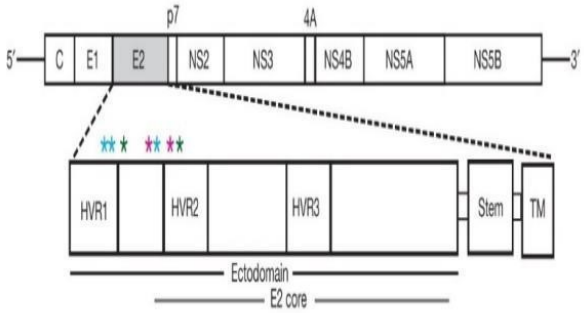

Figure 8: Organization of HCV genome and E2 protein. The HCV genome demonstrates the capsid protein (C), E1 protein and the structural proteins (NS2, NS3, NS4B, NS5A and NS5B). The grey bar indicates E2 core structure while black [33]

The structural analysis of E2 core reveals its packed globular structure with eight disulfide bonds and is highly glycosylated. The secondary structure of E2 core primarily comprises of β sheets and random coils with two small α helix [34].

The E2 core structure exhibits a sandwich of two anti-parallel β sheets (A and B) present at the center. These both front and back anti-parallel sheets are held together by disulfide bonds and intensively hydrophobic core. Sheet A undergoes an IgG like fold while the B sheet constitutes a novel fold [35].

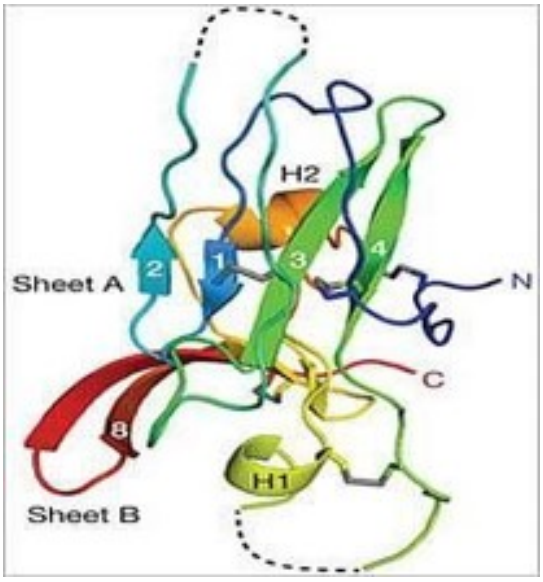

Figure 9: Representation of E2 core domain exhibiting the folding of sheets A and B [33]

11. Hyper Variable Region

The unmitigated genetic diversity of E2 proteins is exhibited through amino acid variations of more than 37% in among various primary isolates of HCV . The E2 protein comprises of three regions of immense variation of amino acids termed as Hyper Variable Regions(HRV). These include HVR1 (384-411), HVR2 (461-481), HVR3 (431-466), and IgVR (570-580) [36]. This diversity attributes to the mechanisms eluding the host immune system. Regardless of this extreme genetic variability in these proteins, there are regions of highly conserved genomic sequence regions exhibiting the N and O linked glycans and high surface glycosylation [37].

HRV1 exhibits the most diverse nature among all hyper variable regions and is located at N terminal ectodomain of E2 with a length of 26-27 amino acids ranging from 384 to 411 [38]. Due to its ability to maintain a certain shape and positively charged amino acid residues, HVR1 is essential for target cell identification and virus attachment [39]. HVR1 serves as a critical mediator in the initial stages of the HCV lifecycle, enhancing viral attachment to hepatocytes. This is largely due to its role in fostering interactions between the virus and various cellular receptors, particularly the scavenger receptor class B type

I (SR-BI) and low-density lipoprotein receptor (LDL-R). Intriguingly, HVR1´s interaction with low density lipo-protein receptor (LDL-R) may involve Apolipoprotein E (ApoE), a component of lipoproteins [40]. Consequently, specific antibodies targeting ApoE can effectively neutralize wild-type HCV particles, making HVR1 a prime significant factor in viral entry and propagation.

HVR1 contains epitopes for neutralizing antibodies and act as a prime target for vaccine formation. Its high sequence variability often compels the virus to mutate, allowing it to dodge the host's immune system [40]. Anti-HVR1 antibodies can disrupt the binding of broadly neutralizing antibodies to HCV, which is a significant aspect of the virus's evasion strategy. The other significant hyper variable region of envelope protein E2 of HCV is HVR 2 comprising of nine amino acids from polyprotein position 705 to 715 and is present down stream the HVR1. While HVR2 may not be the primary target for neutralizing antibodies, it plays a prime role in HCV's entry into host cells. It acts as alliance with HVR1 to influence interactions with key cellular receptors, such as CD81, which play significant role in the viral entry process [41]. The deletion of HVR2 or IgVR regions of E2 protein in HCV disrupts the formation of heterodimers between E1 and E2, which are indispensable for the virus's entry into host cells. This deletion significantly reduces HCV's binding to CD81, a critical host receptor required for successful entry [42]. Another important region exhibiting the high sequence variability in E2 protein is Immunoglobulin Variable Region also known as IgVR [43]. This is located further downstream of the HVR2 near the transmembrane domain of E2. IgVR complements HVR1 and HVR2 in the context of E2, further contributing to the HCV entry process. Similar to HVR2, IgVR does not seem to be the primary target for neutralizing antibodies. However, its absence, when deleted along with HVR2, disrupts the formation of E1E2 heterodimers, thereby affecting viral entry. Certain antibodies raised against the cytosolic portion of E2 exhibit increased reactivity towards E2. These antibodies display heightened effectiveness in reducing the interactions between E2 and CD81 when any of the hypervariable regions are deleted [40]. Of particular interest, some of these antibodies demonstrate remarkable broad neutralization capabilities across all HCV genotypes and are capable of recognizing epitope I. Deleting two or more hypervariable regions enhances their proficiency in inhibiting E2-CD81 interactions.

12. Functions of E2

The HCV E2 protein interacts with the other Envelope glycoprotein E1 to form a heterodimer, which is the main protein that helps the virus enter cells and fuse with other viruses. These two proteins are stabilized by disulfide connections and form bigger units with covalent bonding [44]. The lipid membrane produced by the host cells is injected into this E1-E2 protein heterodimer, which together with other components makes up the HCV envelope.

For cell surface molecules like CD81 and SR-BI, E2 protein serves as the primary receptor binding site [39]. The interaction of the Cd81 receptor binding site with the discontinuous surrounding surface of the three highly conserved areas of the HCV E2 protein serves as a metaphor for the binding reaction of E2 protein with Cd81 cell surface receptors [42]. The interaction between the SR-BI receptors and the N terminal hypervariable regions (HRV1) of the E2 protein is assumed to provide the basis for the binding response of E2 protein with the SR-BI [36].

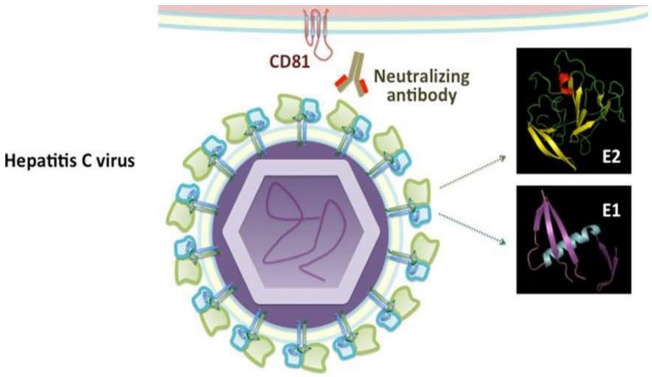

Figure 10: Model of the hepatitis C virus that shows how the E1E2 heterodimers are embedded in the lipid layer on the envelope [41]

13. Role of E1 and E2 in Development of Vaccine Against HCV

Various researches suggest E2 region as the most suitable target for developing a vaccine against HCV because it initiates humoral and cell mediated responses in humans [39]. However due to intensive mutating and genetically varying nature of HCV this task becomes tricky and remains unachievable. E1 and E2 play crucial roles in determining viral pathogenicity and influencing the host immune response, highlighting the significant potential of envelope glycoproteins in the development of a vaccine. Notably, the administration of a prime-boost regimen involving E1/E2 has been shown to induce cross-reactive immune responses in the chimpanzee model, resulting in the production of neutralizing antibodies. This finding suggests the presence of conserved immunogenic epitopes across different genotypes. According to recent studies, the administration of E1/E2 recombinant proteins to human volunteers has yielded the production of cross-neutralizing antibodies capable of targeting all significant genotypes. This outcome serves as a foundation for the advancement of a vaccine, utilizing the recombinant glycoproteins. Since the E2 protein region of HCV contains the most genetically diverse and varying sequences of amino acids, therefore the studies related to this protein will contribute to vaccine development [41]. E2 gains significant attention as the more immunogenic of the two envelope proteins. This is evident from the wealth of monoclonal antibodies raised against E2. In contrast, E1 is relatively less immunogenic during natural infections. Studies of the monoclonal antibodies targeting E1 are less common. However, experiments involving

E1 immunization do generate antibodies [45]. Chronic HCV infections may lead to the production of antibodies recognizing a conserved epitope within the E1 protein, specifically in the amino acid region between aa313-327 [46]. While prior vaccine candidates had limited success, future prospects hinge on eliciting broadly neutralizing antibodies (bnAbs) and an efficient T-cell response. Notably, bnAbs, specifically targeting conserved regions on E1 and E2, offer a promising avenue for vaccine design [47]. However, the structural flexibility of E1 and E2 epitopes presents challenges in vaccine development, necessitating a detailed understanding of their conformational changes during infection [48]. Significant strides have been made in the creation of a vaccine thanks to the HCV glycoproteins' architecture. However, the utilization of E2 alone for vaccine or medication design creates issues given that the majority of the E2 protein is made up of loops and flexible regions. A crucial neutralizing immunodominant face of the protein, the area between HVR-1 and HVR-2, which contains a portion of the AR3C epitope, appears to have substantial structural flexibility. The structural data for the broadly neutralizing AR3C antibody, however, is quite helpful. In addition, because the E2 core keeps its natural fold even after the HVR-1 and HVR-2 are removed, it's possible that future vaccine candidates will forgo including variable regions while still keeping the natural fold.

14. Phylogenetic Analysis

With the exception of Balochistan, where the most prevalent subtype is 1a, frequency research in 2013 showed that the most prevalent HCV genotype in Pakistan is 3a [49].The HCV glycoproteins (E1 and E2), which are hypervariable transmembranes found on the surface of the virus, are among its structural proteins and are crucial for the virus's attachment to the host cell via cell receptors. The C terminal sphere of envelope protein E1 is implicated in changes to membrane permeability and membrane association.

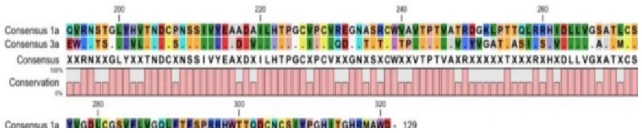

Figure 11: The multiple sequence alignment links the HCV E1 genotype 3a and 1a segregated from different parts of the world to universal consensus sequence. The cryptic alignment symbols are represented by (X), whereas the identical leftovers are shown as (.) [49]

Up to 11 N-Linked glycosylation sites are present on envelope protein E2, which participates in viral entry by interacting with human CD81's extracellular loop, the scavenger receptor class B type 1 /SRB-1), the high viscosity lipoprotein (HDL) binding molecule and the mannose binding proteins DC-SIGN and L-SIGN. Although both glycoproteins are crucial for viral entry, designing vaccines or inhibitory compounds against them is difficult due to their hyper variability. in order to identify conderved peptides in HCV that could serve as valuable targets in the creation of new inhibitory composites and lower the risk of HCV in Pakistan, the study was done to perform a sequence analysis of E1 and E2 among genotypes 3a and 1a. A technique that was grounded in consensus was employed to create efficient peptides. For the HCV E1 (Figure 11) and E2 (Figure 12) proteins, a global consensus sequence was created utilizing the multiple alignment point of the CLC workbench. Recaptured from the NCBI protein database, the HCV E1 and E2 sequences of genotypes 3a and 1a were then subjected to conservation analysis utilizing the multiple sequence alignment feature of the CLC workbench. small peptides of 8 to 25 amino acids were created from the largely conserved residues of E1, and using the same criteria, small peptides of the largely conserved portions of E2 were also created. Given that the 1a and 3a genotypes share these peptides, they may be beneficial for creating peptide-based vaccinations and inhibitory composites.

Figure 12: The multiple sequence alignment links the HCV E1 genotype 3a and 1a segregated from different parts of the world to the universal consensus sequence. The cryptic alignment symbols are represented by (X), whereas the identical leftovers are shown as (.) [49]

Both proteins' phylogenetic trees revealed groupings built on the basis of evolutionary relatedness. It can be concluded that the envelope proteins of genotype 3a, which are geographically isolated, are connected to genotype 1a from many nations and have a common evolutionary ancestor. Additionally, it is implied that a common strain exists between genotypes 1a and 3a. Thus, it can be said that HCV E1 and E2 are related to genotype 1a and 3a in other nations in terms of evolution. Figure 13 displays the evolutionary tree of the E1 protein, while Figure 14 displays the evolutionary tree of the HCV E2 protein.

From the NCBI protein database, 150 HCV E1 and E2 sequences belonging to genotypes 3a and 1a were retrieved. The multiple sequence analysis criterion of the CLC workbench was used to create consensus sequences of E1 and E2 proteins for each genotype. These consensus sequences were then utilized to create E1 and E2 global consensus sequences. Except for Baluchistan, where genotype 1a is the most recent genotype, genotype 3a is reported as the most recent genotype in Pakistan. The goal of this work was to find possibly conserved peptides in the HCV E1 and E2 hypervariable proteins. Largely conserved and variable regions were identified through conservation analysis, and from the variable regions, peptides were created that could potentially be used as targets for the creation of new inhibitory composites that are effective against both the 3a and 1a genotypes, protecting the Pakistani population from the threat

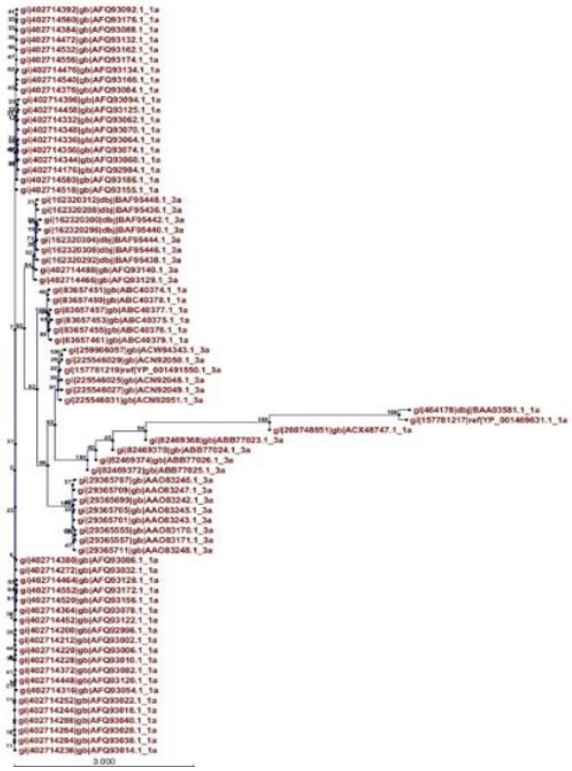

Figure 13: Phylogenetic Tree displaying Evolutionary links between HCV E1 proteins from genotypes 3a and 1a from various geographical locations [49]

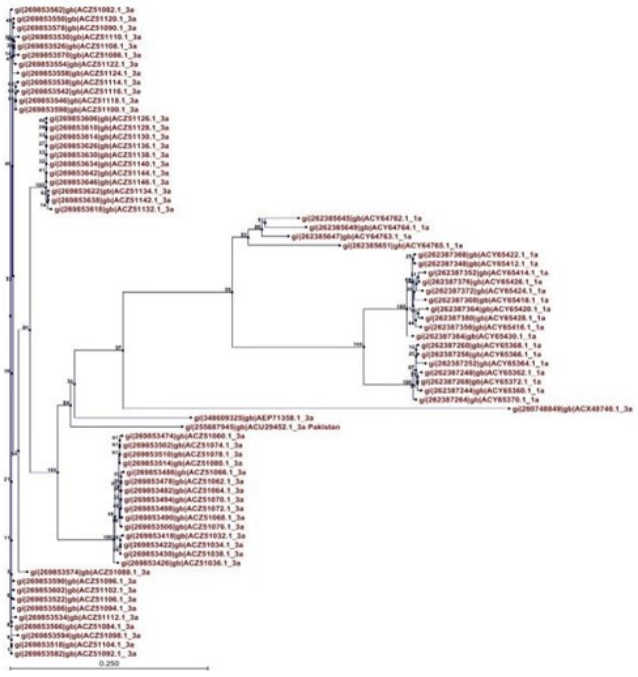

Figure 14: Phylogenetic Tree displaying Evolutionary links between HCV E1 proteins of genotypes 3a and 1a from various geographical Locations [49]

of HCV infection. The evolutionary relationships between various genetic variations or strains of the virus are shown visually in a phylogenetic tree of the HCV envelope protein E2. The evolutionary tree for the Hepatitis C virus´ E2 envelope protein is displayed below

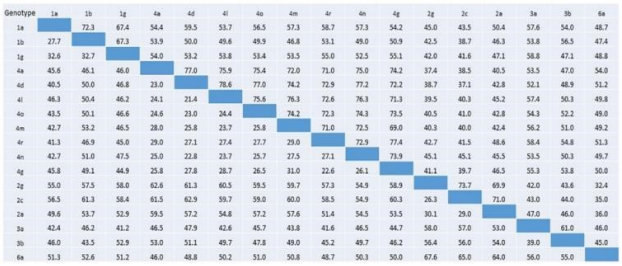

Figure 15: Analysis of HCV E1 protein's percentage nucleotide identity and divergence using the pairwise distance maximum likelihood technique via MEGA 7 [50]

A phylogenetic study was completed in 2019 using 278 sequences at random, representing all genotypes, historical periods, and isolation countries. Distance-based neighbor- joining was used to build the phylogenetic tree with MEGA® V.7.0 software (1,000 replication bootstrap values) [50]. To create a phylogenetic tree, reference genotype sequences were obtained from the https://hcv.lanl.gov database. In order to analyze E1 nucleotide mutations and amino acid changes, all obtained sequences were further pairwise aligned with genotype-specific reference sequences using the MAFFT program. The identification of nucleotide mutations and relative residue analysis of all the sequences were performed using the BioEdit® V.5.0.6(17) program. A different and genotype- specific distribution pattern was shown by the phylogenetic study of the HCV E1 gene. All E1 sequences came together in various clades representing their various genotypes. Between genotypes 1a and 1b, 2c and 2g, and 4a, 4d, and 4o, there is a higher level of nucleotide identity in the E1 residue, according to the pairwise analysis of inheritable distances (Figure 15).

Except for one sequence that was connected to both the 5a and 6a genotypes that belonged to Algeria and Iran separately, the majority of the retrieved sequences belonged to genotypes 1, 2, 3, and 4. A total of 19 genotypes (subtypes) were built in the phylogenetic tree. The phylogenetic analysis revealed the E1 protein's evolutionary dynamics for the dispersed distribution of different HCV genotypes and subtypes in the MENA area. In this respect, the genotype 4a was the most prevalent, followed by 3a and 2c. The highest genotype and subtype diversity was found in South Arabia (G- 1(1a, 1b, 1g), G- 2(2a, 2c), G- 3(3a), and G- 4(4a, 4d, 4n, 4o, 4r, 4s)). Egypt (G- 1(1b, 1g) and G- 4(4a, 4l, 4n, 4m, 4u), Iran (G- 1(1b) and G- 3.

References

[1] N. A. Z. M. Amin, T. N. A. M. Jusoh, A. A. Irekeola, R. H. Shueb, Hepatitis c: A review on current and emerging genotyping assays., Malaysian Journal of Medicine & Health Sciences 19 (5).

[2] L. Puchades Renau, M. Berenguer, Introduction to hepatitis c virus infection: Overview and history of hepatitis c virus therapies, Hemodialysis International 22 (2018) S8–S21.

[3] B. Maasoumy, H. Wedemeyer, Natural history of acute and chronic hepatitis c, Best practice & research Clinical gastroenterology 26 (4) (2012) 401–412.

[4] S. K. Sarin, M. Kumar, Natural history of hcv infection, Hepatology international 6 (2012) 684–695.

[5] H. Basit, I. Tyagi, J. Koirala, Hepatitis, c. statpearls, in: Hepatitis CStatPearls, StatPearls Publishing Treasure Island (FL), 2023.

[6] N. Salari, M. Kazeminia, N. Hemati, M. Ammari-Allahyari, M. Mohammadi, S. Shohaimi, Global prevalence of hepatitis c in general population: A systematic review and meta-analysis, Travel medicine and infectious disease 46 (2022) 102255.

[7] J. Yang, J.-L. Qi, X.-X. Wang, X.-H. Li, R. Jin, B.-Y. Liu, H.-X. Liu, H.-Y. Rao, The burden of hepatitis c virus in the world, china, india, and the united states from 1990 to 2019, Frontiers in Public Health 11 (2023) 1041201.

[8] U. Saleem, N. Aslam, R. Siddique, S. Iqbal, M. Manan, Hepatitis c virus: Its prevalence, risk factors and genotype distribution in pakistan, European Journal of Inflammation 20 (2022) 1721727X221144391.

[9] R. Z. A. Khan, R. S. A. Khan, U. A. Khan, L. Alam, S. R. A. Naqvi, S. Z. A. Naqvi, Chronic hcv infection-comparison of awareness regarding mode of transmission in educated and uneducated patients presenting in gastroenterology/medical clinics in various hospitals of sindh province, Pakistan Armed Forces Medical Journal 72 (5) (2022) 1739.

[10] J.-H. Kao, D.-S. Chen, Transmission of hepatitis c virus in asia: past and present perspectives, Journal of gastroenterology and hepatology 15 (2000) E91–E96.

[11] M. J. Alter, Hcv routes of transmission: what goes around comes around, in: Seminars in liver disease, Vol. 31, © Thieme Medical Publishers, 2011, pp. 340–346.

[12] T. Pietschmann, R. J. Brown, Hepatitis c virus, Trends in microbiology 27 (4) (2019) 379–380.

[13] H. Wan, R. L. Adams, B. D. Lindenbach, A. M. Pyle, The in vivo and in vitro architecture of the hepatitis c virus rna genome uncovers functional rna secondary and tertiary structures, Journal of Virology 96 (8) (2022) e01946–21.

[14] M. T. Deniz, S. Akhan, Hepatitis c virus structure and diagnostic methods, in: Hepatitis C-Recent Advances, IntechOpen, 2023.

[15] G. Dupré, R. Volmer, Influence of viral genome properties on polymerase fidelity, Trends in Genetics 39 (1) (2023) 9–14.

[16] S. Faiz, M. Irfan, S. Farooq, I. A. Khan, H. Iqbal, A.-t. Wahab, M. Shakeel, P. Gong, T. Iftner, M. I. Choudhary, Study of drug resistance-associated genetic mutations, and phylo-genetic analysis of hcv in the province of sindh, pakistan, Scientific Reports 13 (1) (2023) 12213.

[17] J. J. Germer, J. N. Mandrekar, J. L. Bendel, P. S. Mitchell, J. D. Yao, Hepatitis c virus genotypes in clinical specimens tested at a national reference testing laboratory in the united states, Journal of clinical microbiology 49 (8) (2011) 3040–3043.

[18] D. Zarębska-Michaluk, Genotype 3-hepatitis c virusÂ´ last line of defense, World Journal of Gastroenterology 27 (11) (2021) 1006.

[19] G. Schnell, R. Tripathi, P. Krishnan, J. Beyer, T. Reisch, M. Irvin, T. Dekhtyar, C. Setze, L. Rodrigues-Jr, K. Alves, et al., Resistance characterization of hepatitis c virus genotype 2 from japanese patients treated with ombitasvir and paritaprevir/ritonavir, Journal of Medical Virology 90 (1) (2018) 109–119.

[20] K. Al Naamani, S. Al Sinani, M. Deschênes, Epidemiology and treatment of hepatitis c genotypes 5 and 6, Canadian Journal of Gastroenterology and Hepatology 27 (1) (2013) e8–e12.

[21] E. Gupta, J. Samal, A. Pandey, G. Singh, H. A. Gupta, R. Agarwal, M. K. Sharma, Treatment response and drug resistance profiling of genotype 6 of hepatitis c virus in hcv/hiv co-infected patients: a pilot study from india, Viruses 14 (5) (2022) 944.

[22] P. Guntipalli, R. Pakala, S. Kumari Gara, F. Ahmed, A. Bhatnagar, M. Endaya Coronel, A. Razzack, A. Solimando, A. Thompson, K. Andrews, et al., Worldwide prevalence, genotype distribution and management of hepatitis c, Acta Gastroenterol Belg 84 (4) (2021) 637–656.

[23] A. Ahmad, M. R. Khan, R. Naz, H. Akbar, Prevalence of hcv in tehsil babozi district swat, khyber-pakhtunkhwa, pakistan, International Journal of Scientific and Engineering Research 12 (3) (2021) 789–797.

[24] P. Falson, B. Bartosch, K. Alsaleh, B. A. Tews, A. Loquet, Y. Ciczora, L. Riva, C. Montigny, C. Montpellier, G. Duverlie, et al., Hepatitis c virus envelope glycoprotein e1 forms trimers at the surface of the virion, Journal of virology 89 (20) (2015) 10333–10346.

[25] D. Sepulveda-Crespo, S. Resino, I. Martinez, Hepatitis c virus vaccine design: focus on the humoral immune response, Journal of Biomedical Science 27 (2020) 1–12.

[26] Y. Tong, X. Chi, W. Yang, J. Zhong, Functional analysis of hepatitis c virus (hcv) envelope protein e1 using a trans-complementation system reveals a dual role of a putative fusion peptide of e1 in both hcv entry and morphogenesis, Journal of Virology 91 (7) (2017) 10–1128.

[27] Y. Tong, D. Lavillette, Q. Li, J. Zhong, Role of hepatitis c virus envelope glycoprotein e1 in virus entry and assembly, Frontiers in immunology 9 (2018) 1411.

[28] G. Vieyres, J. Dubuisson, T. Pietschmann, Incorporation of hepatitis c virus e1 and e2 glycoproteins: the keystones on a peculiar virion, Viruses 6 (3) (2014) 1149–1187.

[29] N. Echeverría, G. Moratorio, J. Cristina, P. Moreno, Hepatitis c virus genetic variability and evolution, World journal of hepatology 7 (6) (2015) 831.

[30] M. R. Oliver, K. Toon, C. B. Lewis, S. Devlin, R. J. Gifford, J. Grove, Evidence of a novel viral membrane fusion mechanism shared by the hepaci, pegi and pestiviruses, bioRxiv (2022) 2022–10.

[31] F. Stoll-Keller, H. Barth, S. Fafi-Kremer, M. B. Zeisel, T. F. Baumert, Development of hepatitis c virus vaccines: challenges and progress, Expert Review of Vaccines 8 (3) (2009) 333–345.

[32] A. Nayak, N. Pattabiraman, N. Fadra, R. Goldman, S. L. Kosakovsky Pond, R. Mazumder, Structure–function analysis of hepatitis c virus envelope glycoproteins e1 and e2, Journal of Biomolecular Structure and Dynamics 33 (8) (2015) 1682–1694.

[33] A. G. Khan, J. Whidby, M. T. Miller, H. Scarborough, A. V. Zatorski, A. Cygan, A. A. Price, S. A. Yost, C. D. Bohannon, J. Jacob, et al., Structure of the core ectodomain of the hepatitis c virus envelope glycoprotein 2, Nature 509 (7500) (2014) 381–384.

[34] P. Neddermann, L. Tomei, C. Steinkühler, P. Gallinari, A. Tramontano, R. De Francesco, The nonstructural proteins of the hepatitis c virus: structure and functions., Biological chemistry 378 (6) (1997) 469–476.

[35] A. Sabahi, S. L. Uprichard, W. C. Wimley, S. Dash, R. F. Garry, Unexpected structural features of the hepatitis c virus envelope protein 2 ectodomain, Journal of virology 88 (18) (2014) 10280–10288.

[36] G. Sautto, A. W. Tarr, N. Mancini, M. Clementi, Structural and antigenic definition of hepatitis c virus e2 glycoprotein epitopes targeted by monoclonal antibodies, Journal of Immunology Research 2013 (1) (2013) 450963.

[37] N. Kato, Y. Ootsuyama, T. Tanaka, M. Nakagawa, T. Nakazawa, K. Muraiso, S. Ohkoshi, M. Hijikata, K. Shimotohno, Marked sequence diversity in the putative envelope proteins of hepatitis c viruses, Virus Research 22 (2) (1992) 107–123.

[38] D. Smith, Evolution of the hypervariable region of hepatitis c virus, Journal of Viral Hepatitis 6 (1999) 41–46.

[39] A. J. Weiner, H. M. Geysen, C. Christopherson, J. E. Hall, T. J. Mason, G. Saracco, F. Bonino, K. Crawford, C. D. Marion, K. A. Crawford, Evidence for immune selection of hepatitis c virus (hcv) putative envelope glycoprotein variants: potential role in chronic hcv infections., Proceedings of the National Academy of Sciences 89 (8) (1992) 3468–3472.

[40] Y. Alhammad, J. Gu, I. Boo, D. Harrison, K. McCaffrey, P. T. Vietheer, S. Edwards, C. Quinn, F. Coulibaly, P. Poumbourios, et al., Monoclonal antibodies directed toward the hepatitis c virus glycoprotein e2 detect antigenic differences modulated by the n-terminal hypervariable region 1 (hvr1), hvr2, and intergenotypic variable region, Journal of virology 89 (24) (2015) 12245–12261.

[41] H. Freedman, M. R. Logan, J. L. M. Law, M. Houghton, Structure and function of the hepatitis c virus envelope glycoproteins e1 and e2: antiviral and vaccine targets, ACS Infectious Diseases 2 (11) (2016) 749–762.

[42] W. P. Hofmann, C. Sarrazin, B. Kronenberger, B. Schönberger, K. Bruch, S. Zeuzem, Mutations within the cd81-binding sites and hypervariable region 2 of the envelope 2 protein: Correlation with treatment response in hepatitis c virus–infected patients, The Journal of infectious diseases 187 (6) (2003) 982–987.

[43] A. Albecka, R. Montserret, T. Krey, A. W. Tarr, E. Diesis, J. K. Ball, V. Descamps, G. Duverlie, F. Rey, F. Penin, et al., Identification of new

functional regions in hepatitis c virus envelope glycoprotein e2, Journal of virology 85 (4) (2011) 1777–1792.

[44] A. Torrents de la Peña, K. Sliepen, L. Eshun-Wilson, M. L. Newby, J. D. Allen, I. Zon, S. Koekkoek, A. Chumbe, M. Crispin, J. Schinkel, et al., Structure of the hepatitis c virus e1e2 glycoprotein complex, Science 378 (6617) (2022) 263–269.

[45] L. J. Ströh, T. Krey, Hcv glycoprotein structure and implications for b-cell vaccine development, International Journal of Molecular Sciences 21 (18) (2020) 6781.

[46] A. G. Khan, M. T. Miller, J. Marcotrigiano, Hcv glycoprotein structures: what to expect from the unexpected, Current opinion in virology 12 (2015) 53–58.

[47] M. C. Metcalf, B. M. Janus, R. Yin, R. Wang, J. D. Guest, E. Pozharski, M. Law, R. A. Mariuzza, E. A. Toth, B. G. Pierce, et al., Structure of engineered hepatitis c virus e1e2 ectodomain in complex with neutralizing antibodies, Nature Communications 14 (1) (2023) 3980.

[48] L. J. Ströh, T. Krey, Structural insights into hepatitis c virus neutralization, Current Opinion in Virology 60 (2023) 101316.

[49] S. Idrees, U. A. Ashfaq, N. Idrees, Development of global consensus sequence of hcv glycoproteins involved in viral entry, Theoretical Biology and Medical Modelling 10 (2013) 1–7.

[50] M. U. Sohail, Y. Hadi Muhamad, A. T. Asma A, Comparative phylogenetic and residue analysis of hepatitis c virus e1 protein from the middle east and north africa region.